

# PARABOLIC WEIGHTED NORM INEQUALITIES AND PARTIAL DIFFERENTIAL EQUATIONS

JUHA KINNUNEN AND OLLI SAARI

**ABSTRACT.** We introduce a class of weights related to the regularity theory of nonlinear parabolic partial differential equations. In particular, we investigate connections of the parabolic Muckenhoupt weights to the parabolic BMO. The parabolic Muckenhoupt weights need not be doubling and they may grow arbitrarily fast in the time variable. Our main result characterizes them through weak and strong type weighted norm inequalities for forward-in-time maximal operators. In addition, we prove a Jones type factorization result for the parabolic Muckenhoupt weights and a Coifman-Rochberg type characterization of the parabolic BMO through maximal functions. Connections and applications to the doubly nonlinear parabolic PDE are also discussed.

## 1. INTRODUCTION

Muckenhoupt's seminal result characterizes weighted norm inequalities for the Hardy-Littlewood maximal operator through the so called  $A_p$  condition

$$\sup_Q \int_Q w \left( \int_Q w^{1-p'} \right)^{p-1} < \infty, \quad 1 < p < \infty.$$

Here the supremum is taken over all cubes  $Q \subset \mathbb{R}^n$ , and  $w \in L^1_{loc}(\mathbb{R}^n)$  is a nonnegative weight. These weights exhibit many properties that are powerful in applications, such as reverse Hölder inequalities, factorization property, and characterizability through BMO, where BMO refers to the functions of bounded mean oscillation. Moreover, the Muckenhoupt weights play a significant role in the theory of Calderón-Zygmund singular integral operators, see [10].

Another important aspect of the Muckenhoupt weights and BMO is that they also arise in the regularity theory of nonlinear PDEs. More precisely, the logarithm of a nonnegative solution to any PDE of the type

$$\operatorname{div}(|\nabla u|^{p-2} \nabla u) = 0, \quad 1 < p < \infty,$$

---

2010 *Mathematics Subject Classification.* 42B25, 42B37, 35K55.

*Key words and phrases.* Parabolic BMO, weighted norm inequalities, parabolic PDE, doubly nonlinear equations, one-sided weight.

The research is supported by the Academy of Finland and the Väisälä Foundation.

belongs to BMO and the solution itself is a Muckenhoupt weight. This was the crucial observation in [23], where Moser proved the celebrated Harnack inequality for nonnegative solutions of such equations.

Even though the theory of the Muckenhoupt weights is well established by now, many questions related to higher dimensional versions of the one-sided Muckenhoupt condition

$$\sup_{x \in \mathbb{R}, h > 0} \frac{1}{h} \int_{x-h}^x w \left( \frac{1}{h} \int_x^{x+h} w^{1-p'} \right)^{p-1} < \infty$$

remain open. This condition was introduced by Sawyer [28] in connection with ergodic theory. Since then these weights and the one-sided maximal functions have been a subject of intense research; see [2], [3], [7], [18], [19], [20], [21], [22] and [28]. In comparison with the classical  $A_p$  weights, the one-sided  $A_p^+$  weights can be quite general. For example, they may grow exponentially, since any increasing function belongs to  $A_p^+$ . It is remarkable that this class of weights still allows for weighted norm inequalities for some special classes of singular integral operators (see [3]), but the methods are limited to the dimension one.

The first extensions to the higher dimensions of the one-sided weights are by Ombrosi [26]. The subsequent research in [4], [9] and [16] contains many significant advances, but even in the plane many of the most important questions, such as getting the full characterization of the strong type weighted norm inequalities for the corresponding maximal functions, have not received satisfactory answers yet.

In this paper, we propose a new approach which enables us to solve many of the previously unreachable problems. In contrast with the earlier attempts, our point of view is related to Moser's work on the parabolic Harnack inequality in [24] and [25]. More precisely, in the regularity theory for the doubly nonlinear parabolic PDEs of the type

$$(1.1) \quad \frac{\partial(|u|^{p-2}u)}{\partial t} - \operatorname{div}(|\nabla u|^{p-2}\nabla u) = 0, \quad 1 < p < \infty,$$

(see [11], [13], [15], [29]), there is a condition (Definition 3.2) that plays a role identical to that of the classical Muckenhoupt condition in the corresponding elliptic theory. Starting from the parabolic Muckenhoupt condition

$$(1.2) \quad \sup_R \int_{R^-} w \left( \int_{R^+} w^{1-q'} \right)^{q-1} < \infty, \quad 1 < q < \infty,$$

where  $R^\pm$  are space time rectangles with a time lag, we create a theory of parabolic weights. Here we use  $q$  to distinguish from  $p$  in the doubly nonlinear equation. Indeed, they are not related to each other.

The time variable scales as the modulus of the space variable raised to the power  $p$  in the geometry natural for (1.1). Consequently, the Euclidean balls and cubes have to be replaced by parabolic rectangles

respecting this scaling in all estimates. In order to generalize the one-sided theory of weighted norm inequalities, it would be sufficient to work with the case  $p = 2$ . However, in view of the connections to non-linear PDEs (see [27] and [14]), we have decided to develop a general theory for  $1 < p < \infty$ . As far as we know, the results in this work are new even for the heat equation with  $p = 2$ . There are no previous studies about weighted norm inequalities with the same optimal relation to solutions of parabolic partial differential equations.

Observe that the theory of parabolic weights contains the classical  $A_p$  theory as a special case. However, the difference between elliptic and parabolic weights is not only a question of switching from cubes to parabolic rectangles. There is an extra challenge in the regularity theory of (1.1) because of the time lag appearing in the estimates. A similar phenomenon also occurs in the harmonic analysis with one-sided weights, and it has been the main obstacle in the previous approaches [4], [9], [16], and [26]. Except for the one-dimensional case, an extra time lag appears in the arguments. Roughly speaking, a parabolic Muckenhoupt condition without a time lag implies boundedness of maximal operators with a time lag. In our approach, both the maximal operator and the Muckenhoupt condition have a time lag. This allows us to prove the necessity and sufficiency of the parabolic Muckenhoupt condition for both weak and strong type weighted norm inequalities of the corresponding maximal function. Our main technical tools are covering arguments related to the work of Ombrosi et al. [26], [9]; parabolic chaining arguments from [27], and a Calderón-Zygmund argument based on a slicing technique.

Starting from the parabolic Muckenhoupt condition (1.2), we build a complete parabolic theory of one-sided weighted norm inequalities and BMO in the multidimensional case. Our main results are a reverse Hölder inequality (Theorem 5.2), strong type characterizations for weighted norm inequalities for a parabolic forward-in-time maximal function (Theorem 5.4), a Jones type factorization result for parabolic Muckenhoupt weights (Theorem 6.3) and a Coifman-Rochberg type characterization of parabolic BMO through maximal functions (Theorem 7.5). In Section 8, we explain in detail the connection between parabolic Muckenhoupt weights and the doubly nonlinear equation. We refer to [1], [8], [13], [24], [25], [27] and [29] for more on parabolic BMO and its applications to PDEs.

## 2. NOTATION

Throughout the paper, the  $n$  first coordinates of  $\mathbb{R}^{n+1}$  will be called *spatial* and the last one *temporal*. The temporal translations will be important in what follows. Given a set  $E \subset \mathbb{R}^{n+1}$  and  $t \in \mathbb{R}$ , we denote

$$E + t := \{e + (0, \dots, 0, t) : e \in E\}.$$

The exponent  $p$ , with  $1 < p < \infty$ , related to the doubly nonlinear equation (1.1) will be a fixed throughout the paper.

Constants  $C$  without subscript will be generic and the dependencies will be clear from the context. We also write  $K \lesssim 1$  for  $K \leq C$  with  $C$  as above. The dependencies can occasionally be indicated by subscripts or parentheses such as  $K = K(n, p) \lesssim_{n,p} 1$ .

A weight will always mean a real valued positive locally integrable function on  $\mathbb{R}^{n+1}$ . Any such function  $w$  defines a measure absolutely continuous with respect to Lebesgue measure, and for any measurable  $E \subset \mathbb{R}^{n+1}$ , we denote

$$w(E) := \int_E w.$$

We often omit mentioning that a set is assumed to be measurable. They are always assumed to be. For a locally integrable function  $f$ , the integral average is denoted as

$$\frac{1}{|E|} \int_E f = \oint_E f = f_E.$$

The positive part of a function  $f$  is  $(f)^+ = (f)_+ = 1_{\{f>0\}}f$  and the negative part  $(f)^- = (f)_- = -1_{\{f<0\}}f$ .

### 3. PARABOLIC MUCKENHOUT WEIGHTS

Before the definition of the parabolic Muckenhoupt weights, we introduce the parabolic space-time rectangles in the natural geometry for the doubly nonlinear equation.

**Definition 3.1.** Let  $Q(x, l) \subset \mathbb{R}^n$  be a cube with center  $x$  and side length  $l$  and sides parallel to the coordinate axes. Let  $p > 1$  and  $\gamma \in [0, 1)$ . We denote

$$R(x, t, l) = Q(x, l) \times (t - l^p, t + l^p)$$

and

$$R^+(\gamma) = Q(x, l) \times (t + \gamma l^p, t + l^p).$$

The set  $R(x, t, l)$  is called a  $(x, t)$ -centered parabolic rectangle with side  $l$ . We define  $R^-(\gamma)$  as the reflection of  $R^+(\gamma)$  with respect to  $\mathbb{R}^n \times \{t\}$ . The shorthand  $R^\pm$  will be used for  $R^\pm(0)$ .

Now we are ready for the definition of the parabolic Muckenhoupt classes. Observe that there is a time lag in the definition for  $\gamma > 0$ .

**Definition 3.2.** Let  $q > 1$  and  $\gamma \in [0, 1)$ . A weight  $w > 0$  belongs to the parabolic Muckenhoupt class  $A_q^+(\gamma)$ , if

$$(3.1) \quad \sup_R \left( \int_{R^-(\gamma)} w \right) \left( \int_{R^+(\gamma)} w^{1-q'} \right)^{q-1} =: [w]_{A_q^+(\gamma)} < \infty.$$

If the condition above is satisfied with the direction of the time axis reversed, we denote  $w \in A_q^-(\gamma)$ . If  $\gamma$  is clear from the context or unimportant, it will be omitted in the notation.

The case  $A_2^+(\gamma)$  occurs in the regularity theory of parabolic equations, see [24] and [29]. Before investigating the properties of parabolic weights, we briefly discuss how they differ from the ones already present in the literature. The weights of [9] and [16] were defined on the plane, and the sets  $R^\pm(\gamma)$  in Definition 3.2 were replaced by two squares that share exactly one corner point. The definition used in [4] is precisely the same as our Definition 3.2 with  $p = 1$  and  $\gamma = 0$ .

An elementary but useful property of the parabolic Muckenhoupt weights is that they can effectively be approximated by bounded weights.

**Proposition 3.3.** *Assume that  $u, v \in A_q^+(\gamma)$ . Then  $f = \min\{u, v\} \in A_q^+(\gamma)$  and*

$$[f]_{A_q^+} \lesssim [u]_{A_q^+} + [v]_{A_q^+}.$$

*The corresponding result holds for  $\max\{u, v\}$  as well.*

*Proof.* A direct computation gives

$$\begin{aligned} & \left( \int_{R^-(\gamma)} f \right) \left( \int_{R^+(\gamma)} f^{1-q'} \right)^{q-1} \\ & \lesssim \left( \int_{R^-(\gamma)} f \right) \left( \frac{1}{|R^+(\gamma)|} \int_{R^+(\gamma) \cap \{u > v\}} f^{1-q'} \right)^{q-1} \\ & \quad + \left( \int_{R^-(\gamma)} f \right) \left( \frac{1}{|R^+(\gamma)|} \int_{R^+(\gamma) \cap \{u \leq v\}} f^{1-q'} \right)^{q-1} \\ & \leq \left( \int_{R^-(\gamma)} v \right) \left( \frac{1}{|R^+(\gamma)|} \int_{R^+(\gamma) \cap \{u > v\}} v^{1-q'} \right)^{q-1} \\ & \quad + \left( \int_{R^-(\gamma)} u \right) \left( \frac{1}{|R^+(\gamma)|} \int_{R^+(\gamma) \cap \{u \leq v\}} u^{1-q'} \right)^{q-1} \\ & \leq [u]_{A_q^+} + [v]_{A_q^+}. \end{aligned}$$

The result for  $\max\{u, v\}$  is proved in a similar manner.  $\square$

**3.1. Properties of parabolic Muckenhoupt weights.** The special role of the time variable makes the parabolic Muckenhoupt weights quite different from the classical ones. For example, the doubling property does not hold, but it can be replaced by a weaker forward-in-time comparison condition. The next proposition is a collection of useful facts about the parabolic Muckenhoupt condition, the most important of which is the property that the value of  $\gamma \in [0, 1)$  does not play as big a role as one might guess. This is crucial in our arguments. The same phenomenon occurs later in connection with the parabolic BMO.

**Proposition 3.4.** *Let  $\gamma \in [0, 1)$ . Then the following properties hold true.*

- (i) *If  $1 < q < r < \infty$ , then  $A_q^+(\gamma) \subset A_r^+(\gamma)$ .*
- (ii) *Let  $\sigma = w^{1-q'}$ . Then  $\sigma$  is in  $A_{q'}^-(\gamma)$  if and only if  $w \in A_q^+(\gamma)$ .*
- (iii) *Let  $w \in A_q^+(\gamma)$ ,  $\sigma = w^{1-q'}$  and  $t > 0$ . Then*

$$\int_{R^-(\gamma)} w \leq C_t \int_{t+R^-(\gamma)} w \quad \text{and} \quad \int_{R^+(\gamma)} \sigma \leq C_t \int_{-t+R^+(\gamma)} \sigma.$$

- (iv) *If  $w \in A_q^+(\gamma)$ , then we may replace  $R^-(\gamma)$  by  $R^-(\gamma) - a$  and  $R^+(\gamma)$  by  $R^+(\gamma) + b$  for any  $a, b \geq 0$  in the definition of the parabolic Muckenhoupt class. The new condition is satisfied with a different constant  $[w]_{A_q^+}$ .*
- (v) *If  $1 > \gamma' > \gamma$ , then  $A_q^+(\gamma) \subset A_q^+(\gamma')$ .*
- (vi) *Let  $w \in A_q^+(\gamma)$ . Then*

$$w(R^-(\gamma)) \leq C \left( \frac{|R^-(\gamma)|}{|S|} \right)^q w(S)$$

*for every  $S \subset R^+(\gamma)$ .*

- (vii) *If  $w \in A_q^+(\gamma)$  with some  $\gamma \in [0, 1)$ , then  $w \in A_q^+(\gamma')$  for all  $\gamma' \in (0, 1)$ .*

*Proof.* First we observe that (i) follows from Hölder's inequality and (ii) is obvious. For the case  $t + R^-(\gamma) = R^+(\gamma)$  the claim (iii) follows from Jensen's inequality. For a general  $t$ , the result follows from subdividing the rectangles  $R^\pm(\gamma)$  into smaller and possibly overlapping subrectangles and applying the result to them. The property (iv) follows directly from (iii), as does (v) from (iv).

For (vi), take  $S \subset R^+(\gamma)$  and let  $f = 1_S$ . Apply the  $A_q^+(\gamma)$  condition to see that

$$\begin{aligned} \left( \frac{|S|}{|R^+(\gamma)|} \right)^q w(R^-(\gamma)) &= (f_{R^+(\gamma)})^q w(R^-(\gamma)) \\ &\leq \left( \int_{R^+(\gamma)} f^q w \right) \left( \int_{R^+(\gamma)} w^{1-q'} \right)^{q/q'} w(R^-(\gamma)) \\ &\leq C w(S). \end{aligned}$$

For the last property (vii), take  $R = Q(x, l) \times (t - l^p, t + l^p)$ . Let  $\gamma \in (0, 1)$  and suppose that  $w \in A_q^+(\gamma)$ . We will prove that the condition  $A_q^+(2^{-1}\gamma)$  is satisfied. We subdivide  $Q$  into  $2^{nk}$  dyadic subcubes  $\{Q_i\}_{i=1}^{2^{nk}}$ . This gives dimensions for the lower halves of parabolic rectangles  $R_i^-(\gamma)$ . For a given  $Q_i$ , we stack a minimal amount of the rectangles  $R_i^-(\gamma)$  so that they almost pairwise disjointly cover  $Q_i \times (t - l^p, t - 2^{-1}\gamma l^p)$ . The number of  $R_i^-(\gamma)$  needed to cover  $Q \times (t - l^p, t - 2^{-1}\gamma l^p)$

is bounded by

$$2^{nk} \cdot \frac{(1 - 2^{-1}\gamma)l^p}{2^{-nkp}(1 - \gamma)l^p} = 2^{nk(p+1)} \frac{2 - \gamma}{2(1 - \gamma)}.$$

Corresponding to each  $Q_i$ , there is a sequence of at most  $2^k - 1$  vectors  $d_j = 2^{-k-1}le_j$  with  $e_j \in \{0, 1\}^n$  such that

$$Q_i + \sum_j d_j = 2^{-k}Q.$$

Next we show how every rectangle  $R_i(\gamma)$  can be transported to the same spatially central position  $2^{-k}Q$  without losing too much information about their measures. By (vi) we have

$$w(R_i^-(\gamma)) \leq C \left( \frac{|R_i^-(\gamma)|}{|S|} \right)^q w(S)$$

for any  $S \subset R_i^+(\gamma)$ . We choose  $S$  such that its projection onto space variables is  $(Q_i + d_1) \cap Q_i$ , and its projection onto time variable has full length  $(1 - \gamma)(2^{-k}l)^p$ . Then

$$w(R_i^-(\gamma)) \leq C_0 w(S) \leq C_0 w(R_i^{1-}(\gamma))$$

where  $R_i^{1-}(\gamma) \supset S$  is  $Q_i + d_1$  spatially and coincides with  $S$  as a temporal projection. The constant  $C_0$  depends on  $n$  and  $q$ .

Next we repeat the argument to obtain a similar estimate for  $R_i^{1-}(\gamma)$  in the place of  $R_i^-(\gamma)$ . We obtain a new rectangle to the right hand side, on which we repeat the argument again. With  $k$  iterations, we reach an inequality

$$w(R_i^-(\gamma)) \leq C_0^{2^k-1} w(R_i^{*-}(\gamma))$$

where  $R_i^{*-}(\gamma)$  is the parabolic box whose projection onto the coordinates corresponding to the space variables is  $2^{-k}Q$ . The infimum of time coordinates of points in  $R_i^{*-}(\gamma)$  equals

$$\inf\{t : (x, t) \in R_i^-\} + (2^k - 1)(1 + \gamma)(2^{-k}l)^p.$$

As  $p > 1$ , the second term in this sum can be made arbitrarily small. In particular, for a large enough  $k$ , we have

$$(2^k - 1)(1 + \gamma)(2^{-k}l)^p \leq 2 \cdot 2^{-k(p-1)}l^p \leq \frac{1}{100}\gamma l^p.$$

In this fashion, we may choose a suitable finite  $k$  and divide the sets  $R^\pm(2^{-1}\gamma)$  to  $N \lesssim_{n,\gamma} 2^{nkp}$  parts  $R_i^\pm(\gamma)$ . They satisfy

$$w(R_i^-(\gamma)) \leq C_0^{2^k-1} w(R_i^{*-}(\gamma))$$

and

$$\sigma(R_i^+(\gamma)) \leq C_0^{2^k-1} \sigma(R_i^{*+}(\gamma)),$$

where all starred rectangles have their projections onto space variables centered at  $2^{-k}Q$ ; they have equal side length  $2^{-kp}l^p$ , and

$$\frac{1}{2}\gamma l^p \leq d(R_i^{*-}(\gamma), R_j^{*+}(\gamma)) < 2l^p$$

for all  $i, j$ . All this can be done by a choice of  $k$  which is uniform for all rectangles.

It follows that

$$\begin{aligned} & \left( \int_{R^-(2^{-1}\gamma)} w \right) \left( \int_{R^+(2^{-1}\gamma)} w^{1-q'} \right)^{q-1} \\ & \lesssim \sum_{i,j=1}^N \left( \int_{R_i^-(\gamma)} w \right) \left( \int_{R_j^+(\gamma)} w^{1-q'} \right)^{q-1} \\ & \lesssim \sum_{i,j=1}^N \left( \int_{R_i^{*-}(\gamma)} w \right) \left( \int_{R_j^{*+}(\gamma)} w^{1-q'} \right)^{q-1} \\ & \lesssim \sum_{i,j=1}^N C = C(n, p, k, \gamma, q, [w]_{A_q^+}(\gamma)), \end{aligned}$$

where in the last inequality we used (iv). Since the estimate is uniform in  $R$ , the claim follows.  $\square$

#### 4. PARABOLIC MAXIMAL OPERATORS

In this section, we will study parabolic forward-in-time maximal operators, which are closely related to the one-sided maximal operators studied in [4], [9] and [16]. The class of weights in [9], originally introduced by Ombrosi [26], characterizes the weak type inequality for the corresponding maximal operator, but the question about the strong type inequality remains open. On the other hand, Lerner and Ombrosi [16] managed to show that the same class of weights supports strong type boundedness for another class of operators with a time lag. For the boundedness of these operators, however, the condition on weights is not necessary. Later the techniques developed by Berkovits [4] showed that a weight condition without a time lag implies boundedness of maximal operators with a time lag. That approach applied to all dimensions. In our case both the maximal operator and the Muckenhoupt condition have a time lag. This approach, together with scaling of parabolic rectangles, allows us to prove both the necessity and sufficiency of the parabolic Muckenhoupt condition for weak and strong type weighted norm inequalities for the maximal function to be defined next.



**Definition 4.1.** Let  $\gamma \in [0, 1)$ . For  $f \in L^1_{loc}(\mathbb{R}^{n+1})$  define the parabolic maximal function

$$M^{\gamma+} f(x, t) = \sup_{R(x, t)} \int_{R^+(\gamma)} |f|,$$

where the supremum is taken over all parabolic rectangles centered at  $(x, t)$ . If  $\gamma = 0$ , it will be omitted in the notation. The operator  $M^{\gamma-}$  is defined analogously.

The necessity of the  $A_q^+$  condition can be proved in a similar manner to its analogue in the classical Muckenhoupt theory, but already here the geometric flexibility of Definition 3.2 simplifies the statement.

**Lemma 4.2.** *Let  $w$  be a weight such that the operator  $M^{\gamma+} : L^q(w) \rightarrow L^{q,\infty}(w)$  is bounded. Then  $w \in A_q^+(\gamma)$ .*

*Proof.* Take  $f > 0$  and choose  $R$  such that  $f_{S^+} > 0$ , where  $S^+ = R^+$  if  $\gamma = 0$ . If  $\gamma > 0$ ,

$$S^+ = R^-(\gamma) + (1 - \gamma)l^p + 2^p\gamma l^p$$

will do. Redefine  $f = \chi_{S^+} f$ . Take a positive  $\lambda < C_\gamma f_{S^+}$ . With a suitably chosen  $C_\gamma$ , we have

$$w(R^-) \leq w(\{x \in \mathbb{R}^{n+1} : M^{\gamma+} f > \lambda\}) \leq \frac{C}{\lambda^q} \int_{R^+} f^q w.$$

The claim follows letting  $\lambda \rightarrow C_\gamma f = C_\gamma(w + \epsilon)^{1-q'}$  and  $\epsilon \rightarrow 0$ , and concluding by argumentation similar to Proposition 3.4.  $\square$

**4.1. Covering lemmas.** The converse claim requires a couple of special covering lemmas. It is not clear whether the main covering lemma in [9] extends to dimensions higher than two. However, in our geometry the halves of parabolic rectangles are indexed along their spatial centers instead of corner points, which was the case in [9]. This fact will be crucial in the proof of Lemma 4.4, and this enables us to obtain results in the multidimensional case as well.

**Lemma 4.3.** *Let  $R_0$  be a parabolic rectangle, and let  $\mathcal{F}$  be a countable collection of parabolic rectangles with dyadic sidelengths such that for each  $i \in \mathbb{Z}$  we have*

$$\sum_{\substack{P \in \mathcal{F} \\ l(P)=2^i}} 1_{P^-} \lesssim 1.$$

*Moreover, assume that  $P^- \not\subseteq R^-$  for all distinct  $P, R \in \mathcal{F}$ . Then*

$$\sum_{P \in \mathcal{G}} |P| \lesssim |R_0|,$$

*where  $\mathcal{G} = \{P \in \mathcal{F} : P^+ \cap R_0^+ \neq \emptyset, |P| < |R_0|\}$ .*

*Proof.* Recall that  $R^\pm = R^\pm(0)$ . We may write  $\mathcal{G} \subset \mathcal{G}_0(R_0) \cup \mathcal{G}_1$ , where

$$\mathcal{G}_0(R) = \{P \in \mathcal{F} : P \cap \partial R^+, |P| < |R|\}$$

and

$$\mathcal{G}_1 = \{P \in \mathcal{F} : P \subset R_0^+, |P| < |R_0|\}.$$

That is, the rectangles having their upper halves in  $R_0^+$  are either contained in it or they meet its boundary. An estimate for  $\mathcal{G}_0(R)$  with an arbitrary parabolic rectangle  $R$  instead of  $R_0$  will be needed, so that we start with it. Let  $P$  be a parabolic rectangle with the spatial side length  $l(P) = 2^{-i}$ . If  $P \cap \partial R^+ \neq \emptyset$ , then  $P \subset A_i$ , where  $A_i$  can be realized as a collection of  $2(n+1)$  rectangles corresponding to each face of  $R$  such that

$$|A_i| \lesssim 2l(R)^n \cdot 2^{-ip} + 2nl(R)^{p+n-1} \cdot 2^{-i}.$$

Now choosing  $k_0 \in \mathbb{Z}$  such that  $2^{-k_0} < l(R) < 2^{-k_0+1}$ , we get, by the bounded overlap,

$$\sum_{P \in \mathcal{G}_0(R)} |P| = \sum_{i=k_0}^{\infty} \sum_{\substack{P \in \mathcal{G}_0(R) \\ l(P)=2^{-i}}} |P| \lesssim \sum_{i=k_0}^{\infty} |A_i| \lesssim |R|.$$

Once the rectangles meeting the boundary are clear, we proceed to  $\mathcal{G}_1$ . The side lengths of rectangles in  $\mathcal{G}_1$  are bounded from above. Hence there is at least one rectangle with the maximal side length. Let  $\Sigma_1$  be the collection of  $R \in \mathcal{G}_1$  with the maximal side length. We continue recursively. Once  $\Sigma_j$  with  $j = 1, \dots, k$  have been chosen, take the rectangles  $R$  with the maximal side length among the rectangles in  $\mathcal{G}_1$  satisfying

$$R^- \cap \bigcup_{P \in \bigcup_{j=1}^k \Sigma_j} P^- = \emptyset.$$

Let them form the collection  $\Sigma_{k+1}$ . Define the limit collection to be

$$\Sigma = \bigcup_j \Sigma_j.$$

Each  $P \in \mathcal{G}_1$  is either in  $\Sigma$  or  $P^-$  meets  $R^-$  with  $R \in \Sigma$  and  $l(P) < l(R)$ . Otherwise  $P$  would have been chosen to  $\Sigma$ . This implies that

$$\sum_{R \in \mathcal{G}_1} |R| \leq \sum_{R \in \mathcal{G}_1 \cap \Sigma} \left( |R| + \sum_{\substack{P \in \mathcal{G}_1 : P^- \cap R^- \neq \emptyset \\ |P| < |R|}} |P| \right).$$

In the second sum, both  $P$  and  $R$  are in  $\mathcal{F}$ , so  $P^- \not\subset R^-$  by assumption. Thus  $P \cap \partial R^- \neq \emptyset$ , and the sum in the parentheses is controlled by a constant multiple of  $|R|$  (by applying the estimate we have for  $\mathcal{G}_0(\tilde{R})$ )

where  $\tilde{R}$  is a parabolic rectangle with upper half  $R^-$ ). The rectangles in each  $\Sigma_j$  have equal side length so that

$$\begin{aligned} \sum_{R \in \mathcal{G}_1} |R| &\lesssim \sum_{R \in \mathcal{G}_1 \cap \Sigma} |R| = \sum_j \sum_{R \in \mathcal{G}_1 \cap \Sigma_j} |R| \\ &\lesssim \sum_j \left| \bigcup_{R \in \Sigma_j} R \right| \leq \left| \bigcup_{R \in \mathcal{G}_1} R \right| \leq |R_0|. \end{aligned}$$

□

The hypothesis of the next lemma correspond to a covering obtained using the parabolic maximal function, and the conclusion provides us with a covering that has bounded overlap. This fact is analogous to the two-dimensional Lemma 3.1 in [9].

**Lemma 4.4.** *Let  $\lambda > 0$ ,  $f \in L^1_{loc}(\mathbb{R}^{n+1})$  be nonnegative, and  $A \subset \mathbb{R}^{n+1}$  a set of finitely many points such that for each  $x \in A$  there is a parabolic rectangle  $R_x$  with dyadic side length satisfying*

$$(4.1) \quad \int_{R_x^+(\gamma)} f \approx \lambda.$$

*Then there is  $\Gamma \subset A$  such that for each  $x \in \Gamma$  there is  $F_x \subset R_x^+(\gamma)$  with*

- (i)  $A \subset \bigcup_{x \in \Gamma} \overline{R_x^-}$ ,
- (ii)

$$\frac{1}{|R_x|} \int_{F_x} f \gtrsim \lambda \quad \text{and} \quad \sum_{x \in \Gamma} 1_{F_x} \lesssim 1.$$

*Proof.* To simplify the notation, we identify the sets  $R_x^-$  with their closures. Their side lengths are denoted by  $l_x$ . Let  $x_1 \in A$  be a point with maximal temporal coordinate. Recursively, choose  $x_{k+1} \in A \setminus \bigcup_{j=1}^k R_{x_j}^-$ . Denote  $\Delta = \{x_i\}_i$ . This is a finite set. Take  $x \in \Delta$  with maximal  $l_x$  and define  $\Gamma_1 = \{x\}$ . Let  $\Gamma_{k+1} = \Gamma_k \cup \{y\}$  where  $R_y^- \not\subset R_x^-$  for all  $x \in \Gamma_k$  and  $l_y$  is maximal among the  $l_y$  satisfying the criterion. By finiteness the process will stop and let  $\Gamma$  be the final collection.

Given  $x, y \in \Gamma$  with  $l_x = l_y =: r$  and  $x \neq y$ , their Euclidean distance satisfies

$$|x - y| \geq \min \left\{ \frac{1}{2}r, r^p \right\}.$$

There is a dimensional constant  $\alpha \in (0, 1)$  such that  $\alpha R_x \cap \alpha R_y = \emptyset$ , and, given  $z \in \mathbb{R}^{n+1}$ , there is a dimensional constant  $\beta > 0$  such that

$$\bigcup_{x \in \Gamma: z \in R_x} R_x \subset R(z, \beta r).$$

Thus

$$(\beta r)^n (2\beta r)^p = |R(z, \beta r)| \geq \sum_{\substack{x \in \Gamma: l_x = r, \\ z \in R_x}} |\alpha R_x| = (\alpha r)^n (2\alpha r)^p \sum_{\substack{x \in \Gamma \\ l_x = r}} 1_{R_x}(z),$$

and consequently

$$(4.2) \quad \sum_{\substack{x \in \Gamma \\ l_x = r}} 1_{R_x} \lesssim 1.$$

Denote

$$\mathcal{G}_x = \{y \in \Gamma : R_x^+(\gamma) \cap R_y^+(\gamma) \neq \emptyset, |R_y| < |R_x|\}.$$

By inequality (4.2) the assumptions of Lemma 4.3 are fulfilled. Hence

$$\sum_{y \in \mathcal{G}_x} |R_y^+(\gamma)| \lesssim |R_x^+(\gamma)|.$$

By (4.1), we have

$$\sum_{y \in \mathcal{G}_x} \int_{R_y^+(\gamma)} f \lesssim \lambda \sum_{y \in \mathcal{G}_x} |R_y^+(\gamma)| \lesssim \lambda |R_x^+(\gamma)| \lesssim \int_{R_x^+(\gamma)} f.$$

Let the constant in this inequality be  $N$ .

Denote  $s := \#\mathcal{G}_x$ . In case  $s \leq 2N$ , we choose  $F_x = R_x^+(\gamma)$ . If  $s > 2N$ , we define

$$E_i^x = \left\{ z \in R_x^+ : \sum_{y \in \Gamma : l_y < l_x} 1_{R_y^+(\gamma)}(z) \geq i \right\}.$$

Thus  $\sum_i 1_{E_i^x}(z)$  counts the points  $y \in \mathcal{G}_x$  whose rectangles contain  $z$ . Hence

$$\begin{aligned} 2N \int_{E_{2N}^x} f &\leq \sum_{i=1}^s \int_{E_i^x} f = \int_{R_x^+(\gamma)} f \sum_{i=1}^s 1_{E_i^x} \\ &\leq \int_{R_x^+(\gamma)} f \sum_{y \in \mathcal{G}_x} 1_{R_y^+(\gamma)} = \sum_{y \in \mathcal{G}_x} \int_{R_y^+(\gamma)} f \leq N \int_{R_x^+(\gamma)} f. \end{aligned}$$

For the set  $F_x = R_x^+(\gamma) \setminus E_{2N}^x$  we have

$$\int_{F_x} f = \int_{R_x^+(\gamma)} f - \int_{E_{2N}^x} f \geq \frac{1}{2} \int_{R_x^+(\gamma)} f \gtrsim \lambda |R_x^+(\gamma)|.$$

It remains to prove the bounded overlap of  $F_x$ . Take  $z \in \bigcap_{i=1}^k F_{x_i}$ . Take  $x_j$  so that  $l_{x_j}$  is maximal among  $l_{x_i}$ ,  $i = 1, \dots, k$ . By (4.2) there are at most  $C_n$  rectangles with this maximal side length that contain  $z$ . Moreover, their subsets  $F_x$  meet at most  $2N$  upper halves of smaller rectangles, so that  $k \leq 2NC_n$ .  $\square$

**4.2. Weak type inequalities.** Now we can proceed to the proof of the weak type inequality. The proof makes use of a covering argument as in [9] adjusted to the present setting.

**Lemma 4.5.** *Let  $q \geq 1$ ,  $w \in A_q^+(\gamma)$  and  $f \in L^q(w)$ . There is a constant  $C = C(n, \gamma, p, w, q)$  such that*

$$w(\{x \in \mathbb{R}^{n+1} : M^{\gamma+} f > \lambda\}) \leq \frac{C}{\lambda^p} \int |f|^p w$$

for every  $\lambda > 0$ .

*Proof.* We first assume that  $f > 0$  is bounded and compactly supported. Since

$$\begin{aligned} M^{\gamma+} f(x) &= \sup_{h>0} \frac{1}{R(x, h, \gamma)^+} \int_{R(x, h, \gamma)^+} f \\ &\lesssim \sup_{i \in \mathbb{Z}} \frac{1}{R(x, 2^i, 2^{-2}\gamma)^+} \int_{R(x, 2^i, 2^{-2}\gamma)^+} f \\ &= \lim_{j \rightarrow -\infty} \sup_{i \in \mathbb{Z}; i > j} \frac{1}{R(x, 2^i, \gamma')^+} \int_{R(x, 2^i, \gamma')^+} f, \end{aligned}$$

it suffices to consider rectangles with dyadic sidelengths bounded from below provided that we use smaller  $\gamma$ , and the claim will follow from monotone convergence. The actual value of  $\gamma$  is not important because of Proposition 3.4. We may assume that  $w$  is bounded from above and from below (see Proposition 3.3).

Moreover, it suffices to estimate  $w(E)$ , where

$$E = \{x \in \mathbb{R}^{n+1} : \lambda < M^{\gamma+} f \leq 2\lambda\}.$$

Once this has been done, we may sum up the estimates to get

$$\begin{aligned} w(\mathbb{R}^{n+1} \cap \{M^{\gamma+} f > \lambda\}) &= \sum_{i=0}^{\infty} w(\mathbb{R}^{n+1} \cap \{2^i \lambda < M^{\gamma+} f \leq 2^{i+1} \lambda\}) \\ &\leq \sum_{i=0}^{\infty} \frac{1}{2^i} \frac{C}{\lambda^p} \int |f|^p w \leq \frac{C}{\lambda^p} \int |f|^p w. \end{aligned}$$

Let  $K \subset E$  be an arbitrary compact subset. Denote the lower bound for the sidelengths of the parabolic rectangles in the basis of the maximal operator by  $\xi < 1$ . For each  $x \in K$  there is dyadic  $l_x > \xi$  such that

$$\int_{R^+(x, l_x, \gamma)} f \approx \lambda.$$

Denote  $R_x := R(x, l_x)$ . Since  $f \in L^1$ , we have

$$|R_x^+(\gamma)| < \frac{1}{\lambda} \int f = C(\lambda, \|f\|_{L^1}) < \infty.$$

Thus  $\sup_{x \in K} l_x < \infty$ . Let  $a = \min w$ . There is  $\epsilon > 0$ , uniform in  $x$ , such that

$$w((1 + \epsilon)R_x^- \setminus R_x^-) \leq a\xi^{n+p} \leq w(R_x^-)$$

and  $w((1 + \epsilon)R_x^-) \leq 2w(R_x^-)$  hold for all  $x \in K$ . By compactness there is a finite collection of balls  $B(x, \xi^p \epsilon / 2)$  to cover  $K$ . Denote the set of

centers by  $A$ , apply Lemma 4.4 to extract the subcollection  $\Gamma$ . Each  $y \in K$  is in  $B(x, \xi^p \epsilon/2)$  with  $x \in A$ . Each  $x \in A$  is in  $R_z^-$  with  $z \in \Gamma$ , so each  $y \in K$  is in  $B(x, \xi^p \epsilon/2) \subset (1 + \epsilon)R_z^-$ . Thus

$$\begin{aligned} w(K) &\leq \sum_{z \in \Gamma} w((1 + \epsilon)R_z^-) \leq 2 \sum_{z \in \Gamma} w(R_z^-) \\ &\leq \frac{C}{\lambda^q} \sum_{z \in \Gamma} w(R_z^-) \left( \frac{1}{|R_z^+(\gamma)|} \int_{F_z} f \right)^q \\ &\leq \frac{C}{\lambda^q} \sum_{z \in \Gamma} \frac{w(R_z^-)}{|R_z^-|} \left( \int_{R_z^+(\gamma)} w^{1-q'} \right)^{q-1} \int_{F_z} f^q w \\ &\leq \frac{C}{\lambda^q} \int f^q w. \end{aligned}$$

In the last inequality we used the  $A_q^+$  condition together with a modified configuration justified in Proposition 3.4, and the bounded overlap of the sets  $F_z$ .  $\square$

Now we are in a position to summarize the first results about the parabolic Muckenhoupt weights. We begin with the weak type characterization for the operator studied in [4]. Along with this result, the definition in [4] leads to all same results in  $\mathbb{R}^{n+1}$  as the other definition from [9] does in  $\mathbb{R}^2$ . The next theorem holds even in the case  $p = 1$ , which is otherwise excluded in this paper.

**Theorem 4.6.** *Let  $w$  be a weight and  $q > 1$ . Then  $w \in A_q^+(\gamma)$  with  $\gamma = 0$  if and only if  $M^+$  is of  $w$ -weighted weak type  $(q, q)$ .*

*Proof.* Combine Lemma 4.2 and Lemma 4.5.  $\square$

The next theorem is the first main result of this paper. Observe that all the parabolic operators  $M^{\gamma+}$  with  $\gamma \in (0, 1)$  have the same class of good weights. This interesting phenomenon seems to be related to the fact that  $p > 1$ .

**Theorem 4.7.** *Let  $w$  be a weight and  $q > 1$ . Then the following conditions are equivalent:*

- (i)  $w \in A_q^+$  for some  $\gamma \in (0, 1)$ ,
- (ii)  $w \in A_q^+$  for all  $\gamma \in (0, 1)$ ,
- (iii) there is  $\gamma \in (0, 1)$  such that the operator  $M^{\gamma+}$  is of weighted weak type  $(q, q)$  with the weight  $w$ ,
- (iv) the operator  $M^{\gamma+}$  is of weighted weak type  $(q, q)$  with the weight  $w$  for all  $\gamma \in (0, 1)$ .

*Proof.* Lemma 4.2, Lemma 4.5 and (vii) of Proposition 3.4.  $\square$

## 5. REVERSE HÖLDER INEQUALITIES

Parabolic reverse Hölder inequalities had already been studied in [4], and they were used to prove sufficiency of the nonlagged Muckenhoupt condition for the lagged strong type inequality. The proof included the classical argument with self-improving properties and interpolation. Our reverse Hölder inequality will lead to an even stronger self-improving property, and this will give us a characterization of the strong type inequality. We will encounter several challenges. For example, our ambient space does not have the usual dyadic structure. In the classical Muckenhoupt theory this would not be a problem, but here the forwarding in time gives new complications. We will first prove an estimate for the level sets, and then we will use it to conclude the reverse Hölder inequality.

**Lemma 5.1.** *Let  $w \in A_q^+(\gamma)$ ,  $\tilde{R}_0 = Q_0 \times (\tau, \tau + \frac{3}{2}l_0^p)$  and  $\hat{R}_0 = Q_0 \times (\tau, \tau + l_0^p)$ . Then there exist  $C = C([w]_{A_q^+(\gamma)}, n, p)$  and  $\beta \in (0, 1)$  such that for every  $\lambda \geq w_{R_0^-}$ , we have*

$$w(\hat{R}_0 \cap \{w > \lambda\}) \leq C\lambda|\tilde{R}_0 \cap \{w > \beta\lambda\}|.$$

*Proof.* We introduce some notation first. For a parabolic rectangle  $R = Q \times (t_0, t_0 + 2l(Q)^p)$ , we define

$$(5.1) \quad \hat{R} = Q \times (t_0, t_0 + l(Q)^p) \quad \text{and}$$

$$(5.2) \quad \check{R} = Q \times (t_0 + (1 + \gamma)l(Q)^p, \frac{3}{2}l(Q)^p).$$

Here  $\gamma \in (0, 1/2)$ , and by Proposition 3.4, we may replace the sets  $R^\pm(\gamma)$  everywhere by the sets  $\hat{R}$  and  $\check{R}$ . Note that  $\hat{R} = R^-$ . The hat is used to emphasize that  $\hat{R}$  and  $\check{R}$  are admissible in the  $A_q^+$  condition, whereas  $R^-$  is used as the set should be interpreted as a part of a parabolic rectangle. For  $\beta \in (0, 1)$ , the condition  $A_q^+(\gamma)$  gives

$$|\check{R} \cap \{w \leq \beta w_{\hat{R}}\}| \leq \beta^{p'-1} \int_{\check{R}} \frac{w^{1-p'}}{w_{\hat{R}}^{1-p'}} \leq (\beta C)^{p'-1} |\check{R}|.$$

Taking  $\alpha \in (0, 1)$ , we may choose  $\beta$  such that

$$(5.3) \quad |\check{R} \cap \{w > \beta w_{\hat{R}}\}| > \alpha |\check{R}|.$$

Let

$$\mathcal{B} = \{Q \times (t - \frac{1}{2}l(Q)^p, t + \frac{1}{2}l(Q)^p) : Q \subset Q_0 \text{ dyadic}, t \in (0, l^p)\}.$$

Here dyadic means dyadic with respect to  $Q_0$ , and hence the collection  $\mathcal{B}$  consists of the lower parts  $\hat{R}$  of spatially dyadic short parabolic rectangles interpreted as metric balls with respect to

$$d((x, t), (x', t')) = \max\{|x - x'|_\infty, C_p|t - t'|^{1/p}\}.$$

Notice that  $(n + 1)$ -dimensional Lebesgue measure is doubling with respect to  $d$ .

We define a noncentered maximal function with respect to  $\mathcal{B}$  as

$$M_{\mathcal{B}}f(x) = \sup_{\{x\} \subset B \in \mathcal{B}} \int_B f,$$

where the supremum is taken over all sets in  $\mathcal{B}$  that contain  $x$ . By the Lebesgue differentiation theorem, we have

$$\widehat{R}_0 \cap \{w > \lambda\} \subset \{M_{\mathcal{B}}(1_{\widehat{R}_0} w) > \lambda\} =: E$$

up to a null set. Next we will construct a Calderón-Zygmund type cover. The idea is to use dyadic structure to deal with spatial coordinates, then separate the scales, and finally conclude, with one-dimensional arguments, the assumptions of Lemma 4.3.

Define the slice  $E_t = E \cap (\mathbb{R}^n \times \{t\})$  for fixed  $t$ . Since  $\lambda \geq w_{\widehat{R}_0}$ , we may find a collection of maximal dyadic cubes  $Q_i^t \times \{t\} \subset E_t$  such that for each  $Q_i$  there is  $B_i^t \in \mathcal{B}$  with

$$B_i^t \cap (Q_0 \times \{t\}) = Q_i^t \quad \text{and} \quad \int_{B_i^t} w > \lambda.$$

Clearly  $\{B_i^t\}_i$  is pairwise disjoint and covers  $E_t$ . Moreover, since  $Q_i^t$  is maximal, the dyadic parent  $\widehat{Q}_i^t$  of  $Q_i^t$  satisfies

$$\int_{\widehat{Q}_i^t \times I} w \leq \lambda$$

for all intervals  $I \ni t$  with  $|I| = l(\widehat{Q}_i^t)^p$  and especially for the ones with  $\widehat{Q}_i^t \times I \supset B_i^t$ . Hence

$$(5.4) \quad \lambda < \int_{B_i^t} w \lesssim \int_{\widehat{Q}_i^t \times I} w \leq \lambda.$$

We gather the collections corresponding to  $t \in (\tau, \tau + l_0^p)$  together, and separate the resulting collection to subcollections as follows:

$$\mathcal{Q} = \{B_i^t : i \in \mathbb{Z}, t \in (0, l^p)\} = \bigcup_{j \in \mathbb{Z}} \mathcal{Q}_j,$$

where  $\mathcal{Q}_j = \{Q \times I \in \mathcal{Q} : |Q| = 2^{-jn}|Q_0|\}$ . Each  $\mathcal{Q}_j$  can be partitioned into subcollections corresponding to different spatial dyadic cubes  $\mathcal{Q}_j = \bigcup_i \mathcal{Q}_{ji}$ . Here

$$\mathcal{Q}_{ji} = \{Q \times I \in \mathcal{Q}_j : Q = Q_i^t, t \in (\tau, \tau + l^p)\}.$$

If needed, we may reindex the Calderón-Zygmund cubes canonically with  $j$  and  $i$  such that  $j$  tells the dyadic generation and  $i$  specifies the cube such that  $Q_{ji}^t = Q_{ji}^{t'}$ . Then

$$\bigcup_{B \in \mathcal{Q}_{ij}} B \cap \bigcup_{B' \in \mathcal{Q}_{i'j}} B' = \emptyset$$

whenever  $i \neq i'$ . Thus we may identify  $\mathcal{Q}_{ji}$  with a collection of intervals and extract a covering subcollection with an overlap bounded by 2.



Hence we get a covering subcollection of  $\mathcal{Q}_j$  with an overlap bounded by 2, and hence a countable covering subcollection of  $\mathcal{Q}$  such that its restriction to any dyadic scale has an overlap bounded by 2. Denote the final collection by  $\mathcal{F}$ . Its elements are interpreted as lower halves of parabolic rectangles, that is, there are parabolic rectangles  $P$  with  $P^- \in \mathcal{F}$ .

Collect the parabolic halves  $P^- \in \mathcal{F}$  with maximal side length to the collection  $\Sigma_1$ . Recursively, if  $\Sigma_k$  is chosen, collect  $P^- \in \mathcal{F}$  with equal maximal size such that

$$P^+ \cap \bigcup_{Q^- \in \bigcup_{i=1}^k \Sigma_i} Q^+ = \emptyset$$

to the collection  $\Sigma_{k+1}$ . The collections  $\Sigma_k$  share no elements, and their internal overlap is bounded by 2. Since each  $A \in \Sigma_k$  has equal size, the bounded overlap is inherited by the collection

$$\Sigma_k^+ := \{A^+ : A^- \in \Sigma_k\}.$$

Moreover, by construction, if  $A^+ \in \Sigma_i^+$  and  $B^+ \in \Sigma_j^+$  with  $i \neq j$  then  $A^+ \cap B^+ = \emptyset$ . Hence

$$\mathcal{F}' := \bigcup_i \Sigma_i$$

is a collection such that

$$\sum_{P^- \in \mathcal{F}'} 1_{P^+} \leq 2.$$

According to (5.4) and Lemma 4.3, we get

$$\begin{aligned} w(E) &\leq \sum_{B \in \mathcal{F}} w(B) \lesssim \sum_{B \in \mathcal{F}} \lambda |B| \\ &\leq \sum_{P^- \in \mathcal{F}'} \left( \lambda |P^-| + \sum_{\substack{B \in \mathcal{F} \\ B^+ \cap P^+ \neq \emptyset \\ |B| < |P|}} \lambda |B| \right) \lesssim \lambda \sum_{P^- \in \mathcal{F}'} |P^+|. \end{aligned}$$

Then

$$\begin{aligned} w(E) &\lesssim_\gamma \lambda \sum_{P^- \in \mathcal{F}'} |\check{P}| \lesssim \sum_{P^- \in \mathcal{F}'} \lambda |\check{P} \cap \{w > \beta \lambda\}| \\ &\leq \int_{\bigcup_{S^- \in \mathcal{F}'} \check{S} \cap \{w > \beta \lambda\}} \sum_{P^- \in \mathcal{F}'} 1_{P^+} \lesssim \lambda |\tilde{R}_0 \cap \{w > \beta \lambda\}|. \end{aligned}$$

□

The fact that the sets in the estimate given by the above lemma are not equal is reflected to the reverse Hölder inequality as a time lag. This phenomenon is unavoidable, and it was noticed already in the one-dimensional case, see for instance [18].

**Theorem 5.2.** *Let  $w \in A_q^+(\gamma)$  with  $\gamma \in (0, 1)$ . Then there exist  $\delta > 0$  and a constant  $C$  independent of  $R$  such that*

$$\left( \int_{R^-(0)} w^{\delta+1} \right)^{1/(1+\delta)} \leq C \int_{R^+(0)} w.$$

Furthermore, there exists  $\epsilon > 0$  such that  $w \in A_{q-\epsilon}^+(\gamma)$ .

*Proof.* We will consider a truncated weight  $w := \min\{w, m\}$  in order to make quantities bounded. At the end, the claim for general weights will follow by passing to the limit as  $m \rightarrow \infty$ . Without loss of generality, we may take  $R^- = Q \times (0, l^p)$ . Define  $\widehat{R}$  and  $\check{R}$  as in the previous lemma (see (5.1) and (5.2)). In addition, let  $\widetilde{R}$  be the convex hull of  $\widehat{R} \cup \check{R}$ .

Let  $E = \{w > w_{R^-}\}$ . By Lemma 5.1

$$\begin{aligned} \int_{R^- \cap E} w^{\delta+1} &= |R^- \cap E| w_{R^-}^{\delta+1} + \delta \int_{w_{R^-}}^{\infty} \lambda^{\delta-1} w(\{R^- \cap \{w > \lambda\}\}) d\lambda \\ &\leq |R^- \cap E| w_{R^-}^{\delta+1} + C\delta \int_{w_{R^-}}^{\infty} \lambda^{\delta-1} |\{R \cap \{w > \beta\lambda\}\}| d\lambda \\ &\leq |R^- \cap E| w_{R^-}^{\delta+1} + C\delta \int_{\widetilde{R} \cap E} w^{\delta+1}, \end{aligned}$$

which implies that

$$\int_{R^- \cap E} w^{\delta+1} \leq \frac{1}{1 - \delta C} \left( |R^- \cap E| w_{R^-}^{\delta+1} + C\delta \int_{\widetilde{R} \setminus (R^- \cap E)} w^{\delta+1} \right).$$

Consequently

$$\begin{aligned} \int_{R^-} w^{\delta+1} &\leq \frac{2 - \delta C}{1 - \delta C} |R^-| w_{R^-}^{\delta+1} + \frac{C\delta}{1 - \delta C} \int_{\widetilde{R} \setminus R^-} w^{\delta+1} \\ (5.5) \quad &= C_0 |R^-| w_{R^-}^{\delta+1} + C_1 \delta \int_{\widetilde{R} \setminus R^-} w^{\delta+1}. \end{aligned}$$

Then we choose  $l_1^p = 2^{-1} l^p$ . We can cover  $Q$  by  $M_{np}$  subcubes  $\{Q_i^1\}_{i=1}^{M_{np}}$  with  $l(Q_i^1) = l_1$ . Their overlap is bounded by  $M_{np}$ , and so is the overlap of the rectangles

$$\{R_i^{1-}\} = Q_i \times (l_1^p, \frac{3}{2} l^p)$$

that cover  $\widetilde{R} \setminus R^-$  and share the dimensions of the original  $R^-$ . Hence we are in position to iterate. The rectangles  $R_{ij}^{(k+1)-}$  are obtained from

$R_i^{k-}$  as  $R_i^{1-}$  were obtained from  $R^- =: R_i^{0-}$ ,  $i = 1, \dots, M_{np}$ . Thus

$$\begin{aligned} \int_{R^-} w^{\delta+1} &\leq C_0 |R^-| w_{R^-}^{\delta+1} + C_1 \delta \sum_{i=1}^{M_{np}} \int_{R_i^1} w^{\delta+1} \\ &\leq \sum_{j=0}^N \left( C_0^{j+1} (C_1 \delta)^j \sum_{i=1}^{M_{np}} |R_i^{j-}| w_{R_i^{j-}}^{\delta+1} \right) + (C_1 \delta M_{np})^N \int_{\bigcup_{i=1}^{M_{np}} \tilde{R}_i^N \setminus R_i^{N-}} w^{\delta+1} \\ &= I + II. \end{aligned}$$

For the inner sum in the first term we have

$$\sum_{i=1}^{M_{np}} |R_i^{j-}| w_{R_i^{j-}}^{\delta+1} \leq \sum_{i=1}^{M_{np}} 2^{-j\delta n} l^{-\delta(n+p)} \left( \int_{R_i^{j-}} w \right)^{\delta+1} \leq 2^{-j\delta n} l^{n+p} M_{np}^{\delta+1} w_R^{\delta+1}.$$

Thus

$$I \leq \left( \int_R w \right)^{1+\delta} C_0 M_{np}^{\delta+1} l^{n+p} \sum_{j=0}^N (C_1 C_0 \delta)^j 2^{-j\delta n},$$

where the series converges as  $N \rightarrow \infty$  if  $\delta$  is small enough. On the other hand, if  $w$  is bounded, it is clear that  $II \rightarrow 0$  as  $N \rightarrow \infty$ . This proves the claim for bounded  $w$ , hence for truncations  $\min\{w, m\}$ , and the general case follows from the monotone convergence theorem as  $m \rightarrow \infty$ . The self improving property of  $A_q^+(\gamma)$  follows from applying the reverse Hölder inequality coming from the  $A_{q'}^-(\gamma)$  condition satisfied by  $w^{1-q'}$  and using Proposition 3.4.  $\square$

**Remark 5.3.** An easy subdivision argument shows that the reverse Hölder inequality can be obtained for any pair  $R, t + R$  where  $t > 0$ . Namely, we can divide  $R$  to arbitrarily small, possibly overlapping, subrectangles. Then we may apply the estimate to them and sum up. This kind of procedure has been carried out explicitly in [4].

Now we are ready to state the analogue of Muckenhoupt's theorem in its complete form. Once it is established, many results familiar from the classical Muckenhoupt theory follow immediately.

**Theorem 5.4.** *Let  $\gamma_i \in (0, 1)$ ,  $i = 1, 2, 3$ . Then the following conditions are equivalent:*

- (i)  $w \in A_q^+(\gamma_1)$ ,
- (ii) the operator  $M^{\gamma_2+}$  is of weighted weak type  $(q, q)$  with the weight  $w$ ,
- (iii) the operator  $M^{\gamma_3+}$  is of weighted strong type  $(q, q)$  with the weight  $w$ .

*Proof.* Equivalence of  $A_q^+$  and weak type follows from Theorem 4.7. Theorem 5.2 gives  $A_{q-\epsilon}^+$ , so (iii) follows from Marcinkiewicz interpolation and the final implication (iii)  $\Rightarrow$  (ii) is clear.  $\square$

6. FACTORIZATION AND  $A_1^+$  WEIGHTS

In contrast with the classical case, it is not clear what is the correct definition of the parabolic Muckenhoupt class  $A_1^+$ . One option is to derive a  $A_1^+$  condition from the weak type  $(1,1)$  inequality for  $M^{\gamma+}$ , and get a condition that coincides with the formal limit of  $A_q^+$  conditions. We propose a slightly different approach and consider the class arising from factorization of the parabolic Muckenhoupt weights and characterization of the parabolic BMO.

**Definition 6.1.** Let  $\gamma \in [0,1)$ . A weight  $w > 0$  is in  $A_1^+(\gamma)$  if for almost every  $z \in \mathbb{R}^{n+1}$ , we have

$$(6.1) \quad M^{\gamma-}w(z) \leq [w]_{A_1^+(\gamma)}w(z).$$

The class  $A_1^-(\gamma)$  is defined by reversing the direction of time.

The following proposition shows that, in some cases, the  $A_1^+$  condition implies the  $A_1$  type condition equivalent to the weak  $(1,1)$  inequality. Moreover, if  $\gamma = 0$ , then the two conditions are equivalent.

**Proposition 6.2.** Let  $w \in A_1^+(\gamma)$  with  $\gamma < 2^{1-p}$ .

(i) For every parabolic rectangle  $R$  it holds that

$$(6.2) \quad \int_{R^-(2^{p-1}\gamma)} w \lesssim_{\gamma, [w]_{A_1^+}} \inf_{z \in R^+(2^{p-1}\gamma)} w(z).$$

(ii) For all  $q > 1$  we have that  $w \in A_q^+$ .

*Proof.* Denote  $\delta = 2^{p-1}\gamma$ . Take a parabolic rectangle  $R_0$ . We see that every  $z \in R_0^+(\delta)$  is a center of a parabolic rectangle with  $R^-(z, \gamma) \supset R_0^-(\delta)$  such that

$$\int_{R^-(\delta)} w \lesssim \int_{R^-(z, \gamma)} w \leq M^{\gamma-}w(z) \lesssim w(z),$$

where the last inequality used (6.1). This proves (i). The second statement (ii) follows from the fact that (6.2) is an increasing limit of  $A_q^+(\gamma)$  conditions, see Proposition 3.2.  $\square$

Now we will state the main result of this section, that is, the factorization theorem for the parabolic Muckenhoupt weights corresponding to the classical results, for example, in [12] and [5].

**Theorem 6.3.** Let  $\delta \in (0,1)$  and  $\gamma \in (0, \delta 2^{1-p})$ . A weight  $w \in A_q^+(\delta)$  if and only if  $w = uv^{1-p}$ , where  $u \in A_1^+(\gamma)$  and  $v \in A_1^-(\gamma)$ .

*Proof.* Let  $u \in A_1^+(\gamma)$ ,  $v \in A_1^-(\gamma)$  and fix a parabolic rectangle  $R$ . By Proposition 6.2, for all  $x \in R^+(\delta)$ , we have

$$u(x)^{-1} \leq \sup_{x \in R^+(\delta)} u(x)^{-1} = \left( \inf_{x \in R^+(\delta)} u(x) \right)^{-1} \lesssim \left( \int_{R^-(\delta)} u \right)^{-1},$$

and, for all  $y \in R^-(\delta)$ , we have the corresponding inequality for  $v$ , that is,

$$v(y)^{-1} \leq \sup_{y \in R^-(\delta)} v(y)^{-1} = \left( \inf_{y \in R^-(\delta)} v(y) \right)^{-1} \lesssim \left( \int_{R^+(\delta)} v \right)^{-1}.$$

Hence

$$\begin{aligned} & \left( \int_{R^-(\delta)} uv^{1-q} \right) \left( \int_{R^+(\delta)} u^{1-q'} v \right)^{p-1} \\ & \lesssim \left( \int_{R^-(\delta)} u \right) \left( \int_{R^+(\delta)} v \right)^{1-q} \left( \int_{R^+(\delta)} v \right)^{q-1} \left( \int_{R^-(\delta)} u \right)^{-1} = C, \end{aligned}$$

which proves that  $uv^{1-q} \in A_q^+(\delta)$ . The finite constant  $C$  depends only on  $\gamma, \delta, [u]_{A_1^+(\gamma)}$  and  $[v]_{A_1^-(\gamma)}$ .

For the other direction, fix  $q \geq 2$  and  $w \in A_q^+$ . Define an operator  $T$  as

$$Tf = (w^{-1/q} M^{\gamma-} (f^{q-1} w^{1/q}))^{1/(q-1)} + w^{1/q} M^{\gamma+} (fw^{-1/q}).$$

By boundedness of the operators

$$M^{\gamma+} : L^q(w) \rightarrow L^q(w) \quad \text{and} \quad M^{\gamma-} : L^{q'}(w^{1-p'}) \rightarrow L^{q'}(w^{1-p'})$$

we conclude that  $T : L^q \rightarrow L^q$  is bounded. Let

$$B(w) := \|T\|_{L^q \rightarrow L^q} \approx_{[w]_{A_q^+}} 1.$$

Take  $f_0 \in L^q$  with  $\|f_0\|_{L^q} = 1$ . Let

$$\phi = \sum_{i=1}^{\infty} (2B(w))^{-i} T^i f_0$$

where  $T^i$  simply means the  $i$ th iterate of  $T$ . We define

$$u = w^{1/q} \phi^{q-1} \quad \text{and} \quad v = w^{-1/q} \phi.$$

Clearly  $w = uv^{1-q}$ . We claim that  $u \in A_1^+$  and  $v \in A_1^-$ . Since  $q \geq 2$  the operator  $T$  is sublinear, and we obtain

$$\begin{aligned} T(\phi) & \leq 2B(w) \sum_{i=1}^{\infty} (2B(w))^{-(i+1)} T^{i+1}(f_0) \\ & = 2B(w) \left( \phi - \frac{T(f_0)}{2B(w)} \right) \leq 2B(w)\phi. \end{aligned}$$

Noting that  $\phi = (w^{-1/q} u)^{1/(q-1)} = w^{1/q} v$  and inserting the above inequality into the definition of  $T$ , we obtain

$$M^{\gamma-} u \leq (2B(w))^{q-1} u \quad \text{and} \quad M^{\gamma+} v \leq 2B(w)v.$$

This implies that  $u \in A_1^+$  and  $v \in A_1^-$  so the proof is complete for  $q \geq 2$ . Once the claim is known for  $q \geq 2$ , the complementary case  $1 < q < 2$  follows from Proposition 3.4 (ii).  $\square$

Next we will characterize  $A_1^+$  weights as small powers of maximal functions up to a multiplication by bounded functions. The following result looks very much like the classical characterization of Muckenhoupt  $A_1$  weights. However, we emphasize that even if the maximal operator  $M^{\gamma+}$  is dominated by the Hardy-Littlewood maximal operator, the assumptions of the following lemma are not restrictive at all when it comes to the measure  $\mu$ . Indeed, the condition  $M^{\gamma-}\mu < \infty$  almost everywhere still includes rather rough measures. For instance, their growth towards the positive time direction can be almost arbitrary, and the same property is carried over to the  $A_1^+$  weights.

**Lemma 6.4.** (i) *Let  $\mu$  be a locally finite nonnegative Borel measure on  $\mathbb{R}^{n+1}$  such that  $M^-\mu < \infty$  almost everywhere. If  $\delta \in [0, 1)$ , then*

$$w := (M^-\mu)^\delta \in A_1^+(0)$$

*with  $[w]_{A_1^+(0)}$  independent of  $\mu$ .*

(ii) *Let  $w \in A_1^+(\gamma')$ . Then there exists a  $\mu$  as above,  $\delta \in [0, 1)$  and  $K$  with  $K, K^{-1} \in L^\infty$  such that*

$$w = K(M^{\gamma-}\mu)^\delta,$$

*where  $\gamma' < \gamma$ .*

*Proof.* Let  $x \in \mathbb{R}^{n+1}$  and fix a parabolic rectangle  $R_0$  centered at  $x$ . Denote  $\tilde{B} = (2R_0)^-$ . Decompose  $\mu = \mu_1 + \mu_2$  where  $\mu_1 = \mu|_{\tilde{B}}$  and  $\mu_2 = \mu|_{\tilde{B}^c}$ . Kolmogorov's inequality gives

$$\int_{R_0^-} (M^-\mu_1)^\delta \leq C|R_0^-|^{-\delta} \mu_1(\tilde{B})^\delta \leq C \left( \frac{\mu(\tilde{B})}{|\tilde{B}|} \right)^\delta \leq CM^-\mu(x)^\delta.$$

On the other hand, for any  $y \in R_0^-$  and a rectangle  $R(y, L) \cap (\tilde{B})^c \neq \emptyset$ , we have  $L \gtrsim l(R_0)$ . Moreover,  $R(y, L) \subset R(x, CL)$  so that

$$M^-\mu_2(y) \lesssim M^-\mu(x)$$

and

$$\int_{R_0^-} (M^-\mu)^\delta \leq \int_{R_0^-} (M^-\mu_2)^\delta + \int_{R_0^-} (M^-\mu_1)^\delta \lesssim M^-\mu(x)^\delta.$$

To prove (ii), take  $w \in A_1^+(\gamma')$  and a parabolic rectangle  $R$  centered at  $x$ . By the reverse Hölder property (Theorem 5.2), Remark 5.3, and inequality (6.1) we have

$$\left( \int_{R^-(\gamma)} w^{1+\epsilon} \right)^{1/(1+\epsilon)} \lesssim w(x).$$

Denote  $\mu = w^{1+\epsilon}$  and  $\delta = 1/(1+\epsilon)$ . By the Lebesgue differentiation theorem

$$w(x) \leq M^{\gamma-}\mu(x)^\delta \lesssim w(x).$$

Hence

$$K = \frac{w}{(M^{\gamma-\mu})^\delta}$$

is bounded from above and from below, which proves the claim.  $\square$

## 7. A CHARACTERIZATION OF THE PARABOLIC BMO

In this section we discuss the connection between parabolic Muckenhoupt weights and the parabolic BMO. The parabolic BMO was explicitly defined by Fabes and Garofalo in [8], who gave a simplified proof of the parabolic John-Nirenberg lemma in Moser's paper [24]. We consider a slightly modified definition in order to make the parabolic BMO a larger space and a more robust class, see [27]. Our definition has essentially the same connections to PDE as the one in [8]. Moreover, this extends the theory beyond the quadratic growth case and applies to the doubly nonlinear parabolic equations.

**Definition 7.1.** A function  $u \in L^1_{loc}(\mathbb{R}^{n+1})$  belongs to  $\text{PBMO}^+$ , if there are constants  $a_R$ , that may depend on the parabolic rectangles  $R$ , such that

$$(7.1) \quad \sup_R \left( \int_{R^+(\gamma)} (u - a_R)^+ + \int_{R^-(\gamma)} (a_R - u)^+ \right) < \infty.$$

for some  $\gamma \in (0, 1)$ . If (7.1) holds with the time axis reversed, then  $u \in \text{PBMO}^-$ .

If (7.1) holds for some  $\gamma \in (0, 1)$ , then it holds for all of them. Moreover, we can consider prolonged parabolic rectangles  $Q \times (t - Tl^p, t + Tl^p)$  with  $T > 0$  and still recover the same class of functions. These facts follow from the main result in [27], and they can be deduced from results in [1] and in a special case from results in [8].

The fact that  $\gamma > 0$  is crucial. For example, the John-Nirenberg inequality (Lemma 7.2) for the parabolic BMO cannot hold without a time lag. Hence a space with  $\gamma = 0$  cannot be characterized through the John-Nirenberg inequality. The following lemma can be found in [27]. See also [8] and [1].

**Lemma 7.2.** *Let  $u \in \text{PBMO}^+$  and  $\gamma \in (0, 1)$ . Then there are  $A, B > 0$  depending only on  $n, \gamma$  and  $u$  such that*

$$(7.2) \quad |R^+(\gamma) \cap \{(u - a_R)^+ > \lambda\}| \leq Ae^{-B\lambda}|R|$$

and

$$(7.3) \quad |R^-(\gamma) \cap \{(a_R - u)^+ > \lambda\}| \leq Ae^{-B\lambda}|R|.$$

There are also more elementary properties that can be seen from the Definition 7.1. Since we will need them later, they will be stated in the next proposition.

**Proposition 7.3.** (i) If  $u, v \in \text{PBMO}^+$  and  $\alpha, \beta \in (0, \infty)$ , then  $\alpha u + \beta v \in \text{PBMO}^+$ .  
(ii)  $u \in \text{PBMO}^+$  if and only if  $-u \in \text{PBMO}^-$ .

*Proof.* For (i), note that

$$(u + v - (a_R^u + a_R^v))^+ \leq (u - a_R^u)^+ + (u - a_R^v)^+,$$

and an analogous estimate holds for the negative part. Hence  $\alpha u + \beta v \in \text{PBMO}^+$  with

$$a_R = \frac{a_R^u}{\alpha} + \frac{a_R^v}{\beta}.$$

Since

$$(u - a_R)^+ = ((-u) - (-a_R))^- \quad \text{and} \quad (u - a_R)^- = ((-u) - (-a_R))^+,$$

the second assertion is clear.  $\square$

The goal of this section is to characterize the parabolic BMO in the sense of Coifman and Rochberg [6]. The Muckenhoupt theory developed so far gives a characterization for the parabolic Muckenhoupt weights, so what remains to do is to prove the equivalence of the parabolic BMO and the  $A_q^+$  condition.

**Lemma 7.4.** Let  $q \in (1, \infty)$  and  $\gamma \in (0, 1)$ . Then

$$(7.4) \quad \text{PBMO}^+ = \{-\lambda \log w : w \in A_q^+(\gamma), \lambda \in (0, \infty)\}.$$

*Proof.* We abbreviate  $R^\pm(\gamma) = R^\pm$  even if  $\gamma \neq 0$ . For  $u \in \text{PBMO}^+$ , Lemma 7.2 gives  $\epsilon > 0$  such that

$$\int_{R^-} e^{-\epsilon u} = e^{-a_R \epsilon} \int_{R^-} e^{\epsilon(a_R - u)} \leq e^{-a_R \epsilon} \int_{R^-} e^{\epsilon(a_R - u)^+} \leq C_- e^{-a_R \epsilon}$$

and, for some  $q < \infty$ ,

$$\begin{aligned} \int_{R^+} e^{\epsilon u/(q-1)} &= e^{a_R \epsilon/(q-1)} \int_{R^+} e^{(u - a_R) \epsilon/(q-1)} \\ &\leq e^{a_R \epsilon/(q-1)} \int_{R^+} e^{(u - a_R)^+ \epsilon/(q-1)} \leq C_+ e^{a_R \epsilon/(q-1)} \end{aligned}$$

so  $w := e^{-u\epsilon} \in A_q^+$  and  $u = -\epsilon^{-1} \cdot \log w$  as it was claimed.

To prove the other direction, take  $w \in A_q^+$  with  $q \leq 2$ . Choose

$$a_R = \log w_{R^-}.$$



Then by Jensen's inequality and the parabolic Muckenhoupt condition, we have

$$\begin{aligned}
\exp \int_{R^+} (a_R - \log w)^+ &\leq \int_{R^+} \exp(a_R - \log w)^+ \\
&\leq 1 + \int_{R^+} \exp \left( a_R - \frac{1}{1-q'} \log w^{1-q'} \right) \\
&\leq 1 + \exp(a_R) \left( \int_{R^+} w^{1-q'} \right)^{q-1} \\
&= 1 + w_{R^-} \left( \int_{R^+} w^{1-q'} \right)^{q-1} \leq 1 + C_{A_q^+}.
\end{aligned}$$

On the other hand, again by Jensen's inequality,

$$\begin{aligned}
\exp \int_{R^-} (\log w - a_R)^+ &\leq \int_{R^-} \exp(\log w - a_R)^+ \\
&\leq 1 + \int_{R^-} \exp(\log w - a_R) \\
&\leq 1 + \exp(-a_R) \int_{R^-} w \\
&\leq 1 + w_{R^-}^{-1} w_{R^-} \leq 2.
\end{aligned}$$

This implies that

$$\begin{aligned}
&\log(2(1 + C_{A_q^+})) \\
&\geq \int_{R^+} (-\log w - (-a_R))^+ + \int_{R^-} (-a_R - (-\log w))^+,
\end{aligned}$$

and  $u = -\log w \in \text{PBMO}^+$ . Applying the same argument for  $A_q^-$  with  $q > 2$  shows that  $-\log w^{1-q'} \in \text{PBMO}^-$  and consequently Proposition 7.3 implies that  $-(q' - 1) \log w \in \text{PBMO}^+$ .  $\square$

The following Coifman-Rochberg [6] type characterization for the parabolic BMO is the main result of this section. Observe, that it gives us a method to construct functions of parabolic bounded mean oscillation with prescribed singularities.

**Theorem 7.5.** *If  $f \in \text{PBMO}^+$  then there exist  $\gamma \in (0, 1)$ , constants  $\alpha, \beta > 0$ , a bounded function  $b \in L^\infty$  and non-negative Borel measures  $\mu$  and  $\nu$  such that*

$$f = -\alpha \log M^{\gamma-} \mu + \beta \log M^{\gamma+} \nu + b.$$

*Conversely, any  $f$  of the form above with  $\gamma = 0$  and  $M^-\mu, M^+\nu < \infty$  belongs to  $\text{PBMO}^+$ .*

*Proof.* Take first  $f \in \text{PBMO}^+$ . By Lemma 7.4

$$f = -C \log w$$

with  $C > 0$  and  $w \in A_2^+$ . By Theorem 6.3, there are  $u \in A_1^+$  and  $v \in A_1^-$  satisfying the corresponding maximal function estimates (6.1) such that

$$w = uv^{-1}.$$

By Lemma 6.4, there exist functions  $K_u, K_v, K_u^{-1}, K_v^{-1} \in L^\infty$  and non-negative Borel measures  $\mu$  and  $\nu$  such that

$$u = K_u(M^{\gamma^-}\mu)^\alpha \quad \text{and} \quad v = K_v(M^{\gamma^+}\nu)^\beta.$$

Hence  $f$  is of the desired form. The other direction follows from Lemma 6.4.  $\square$

## 8. DOUBLY NONLINEAR EQUATION

We begin with pointing out that the theory discussed here applies not only to (1.1) but also to the PDEs

$$\frac{\partial(|u|^{p-2}u)}{\partial t} - \operatorname{div} A(x, t, u, Du) = 0, \quad 1 < p < \infty,$$

where  $A$  satisfies the growth conditions

$$A(x, t, u, Du) \cdot Du \geq C_0|Du|^p$$

and

$$|A(x, t, u, Du)| \leq C_1|Du|^{p-1}.$$

See [13] and [27] for more. For simplicity, we have chosen to focus on the prototype equation (1.1) here.

**8.1. Supersolutions are weights.** We say that

$$v \in L_{loc}^p((-\infty, \infty); W_{loc}^{1,p}(\mathbb{R}^{n+1}))$$

is a supersolution to (1.1) provided

$$\int \left( |\nabla v|^{p-2} \nabla v \cdot \nabla \phi - |v|^{p-2} v \frac{\partial \phi}{\partial t} \right) \geq 0$$

for all non-negative  $\phi \in C_0^\infty(\mathbb{R}^{n+1})$ . If the reversed inequality is satisfied, we call  $u$  a subsolution. If a function is both sub- and supersolution, it is a weak solution.

The definition above allows us to use the following a priori estimate, which is Lemma 6.1 in [13]. Similar results can also be found in [24] and [29], but we emphasize that the following lemma applies to the full range  $1 < p < \infty$  instead of just  $p = 2$ .

**Lemma 8.1** (Kinnunen–Kuusi [13]). *Suppose that  $v > 0$  is a supersolution of the doubly nonlinear equation in  $\sigma R$  where  $\sigma > 1$  and  $R$  is a parabolic rectangle. Then there are constants  $C = C(p, \sigma, n)$ ,  $C' = C'(p, \sigma, n)$  and  $\beta = \beta(R)$  such that*

$$|R^- \cap \{\log v > \lambda + \beta + C'\}| \leq \frac{C}{\lambda^{p-1}} |R^-|$$

and

$$|R^+ \cap \{\log v < -\lambda + \beta - C'\}| \leq \frac{C}{\lambda^{p-1}} |R^+|$$

for all  $\lambda > 0$ .

**Remark 8.2.** There is a technical assumption  $v > \rho > 0$  in [13]. However, this assumption can be removed, see [17]. Indeed, Lemma 2.3 of [17] improves the inequality (3.1) of [13] as to make the proof of the above lemma work with general  $v > 0$  in the case of (1.1) or more general parabolic quasiminimizers.

Let  $v$  be a positive supersolution and set  $u = -\log v$ . We apply Lemma 8.1 together with Cavalieri's principle to obtain

$$\int_{R^+} (u - a_R)_+^b + \int_{R^-} (a_R - u)_+^b < C(p, \sigma, \gamma, n)$$

with  $b = \min\{(p-1)/2, 1\}$ . A general form of the John-Nirenberg inequality from [1] together with its local-to-global properties from [27] can be used to obtain

$$\int_{R^+(\gamma)} (u - a_R)_+ + \int_{R^-(\gamma)} (a_R - u)_+ < C(p, \sigma, \gamma, n).$$

Hence  $u = -\log v$  belongs to  $\text{PBMO}^+$  in the sense of Definition 7.1. The computations required in this passage are carried out in detail in Lemma 6.3 of [27]. We collect the results into the following proposition, whose content, up to notation, is folklore by now.

**Proposition 8.3.** *Let  $v > 0$  be a supersolution to (1.1) in  $\mathbb{R}^{n+1}$ . Then*

$$u = -\log v \in \text{PBMO}^+$$

*In addition,  $v \in \cap_{q>1} A_q^+$ .*

**Remark 8.4.** This gives a way to construct nontrivial examples of the parabolic Muckenhoupt weights and parabolic BMO functions.

Since  $\log v \in \text{PBMO}^-$ , we have that some power of the positive supersolution  $w$  satisfies a local  $A_2^+(\gamma)$  condition. This follows from Lemma 7.4. However, working a bit more with the PDE, it is possible to prove a weak Harnack estimate which implies the improved weight condition stated in the above proposition. This has been done in [13], but the refinement provided in [17] is again needed in order to cover all positive supersolutions.

**8.2. Applications.** The previous proposition asserts that the definitions of parabolic weights and parabolic BMO are correct from the point of view of doubly nonlinear equation. These properties can be used to deduce two interesting results, the second one of which is new. The first one is a global integrability result for supersolutions, see Theorem 6.5 from [27]. The second application of the parabolic theory of weights is related to singularities of supersolutions. It follows from

Proposition 8.3 and Theorem 7.5. In qualitative terms, the following theorem tells quite explicitly what kind of functions the generic positive supersolutions are.

**Theorem 8.5.** *Let  $v > 0$  be a supersolution to (1.1) in  $\mathbb{R}^{n+1}$ . Then there are positive Borel measures  $\nu$  and  $\mu$  with*

$$M^{\gamma^-}\nu < \infty \quad \text{and} \quad M^{\gamma^+}\mu < \infty,$$

*numbers  $\alpha, \beta > 0$ , and a positive function  $b$  with  $b, b^{-1} \in L^\infty(\mathbb{R}^{n+1})$  so that*

$$v = b \frac{(M^{\gamma^-}\nu)^\alpha}{(M^{\gamma^+}\mu)^\beta}.$$

## REFERENCES

- [1] H. Aimar, *Elliptic and parabolic BMO and Harnack's inequality*, Trans. Amer. Math. Soc. **306** (1988), 265–276.
- [2] H. Aimar and R. Crescimbeni, *On one-sided BMO and Lipschitz functions*, Ann. Sc. Norm. Sup. Pisa **27** (1998), 437–456.
- [3] H. Aimar, L. Forzani and F.J. Martín-Reyes, *On weighted inequalities for singular integrals*, Proc. Amer. Math. Soc. **125** (1997), 2057–2064.
- [4] L. Berkovits, *Parabolic Muckenhoupt weights in the Euclidean space*, J. Math. Anal. Appl. **379** (2011), 524–537.
- [5] R. Coifman, P.W. Jones and J.L. Rubio de Francia, *Constructive decomposition of BMO functions and factorization of  $A_p$  weights*, Proc. Amer. Math. Soc. **87** (1983), 675–676.
- [6] R. Coifman and R. Rochberg, *Another characterization of BMO*, Proc. Amer. Math. Soc. **79** (1980), 249–254.
- [7] D. Cruz-Uribe SFO, C.J. Neugebauer and V. Olesen, *The one-sided minimal operator and the one-sided reverse Hölder inequality*, Studia Math. **116** (1995), 255–270.
- [8] E.B. Fabes and N. Garofalo, *Parabolic B.M.O. and Harnack's inequality*, Proc. Amer. Math. Soc. **95** (1985), 63–69.
- [9] L. Forzani, F.J. Martín-Reyes and S. Ombrosi, *Weighted inequalities for the two-dimensional one-sided Hardy-Littlewood maximal function*, Trans. Amer. Math. Soc. **363** (2011), 1699–1719.
- [10] J. García-Cuerva and J.L. Rubio de Francia, *Weighted Norm Inequalities and Related Topics*, North Holland, Amsterdam, 1985.
- [11] U. Gianazza and V. Vespri, *A Harnack inequality for solutions of doubly nonlinear parabolic equations*, J. Appl. Funct. Anal. **1** (2006), 271–284.
- [12] P.W. Jones, *Factorization of  $A_p$  weights*, Ann. of Math. **111** (1980), 511–530.
- [13] J. Kinnunen and T. Kuusi, *Local behaviour of solutions to doubly nonlinear parabolic equations*, Math. Ann. **337** (2007), 705–728.
- [14] J. Kinnunen and O. Saari, *On weights satisfying parabolic Muckenhoupt conditions*, Nonlinear Anal. **131** (2016), 289–299.
- [15] T. Kuusi, J. Siljander and J.M. Urbano, *Local Hölder continuity for doubly nonlinear parabolic equations*, Indiana Univ. Math. J. **61** (2012), 399–430.
- [16] A. Lerner and S. Ombrosi, *A boundedness criterion for general maximal operators*, Publ. Mat. **54** (2010), 53–71.
- [17] P.-A. Ivert, N. Marola and M. Masson, *Energy estimates for variational minimizers of a doubly nonlinear equation on metric measure spaces*, Ann. Acad. Sci. Fenn. Math. **39** (2014), 711–719.

- [18] F.J. Martín-Reyes, *New proofs of weighted inequalities for the one-sided Hardy-Littlewood maximal functions*, Proc. Amer. Math. Soc. **117** (1993), 691–698.
- [19] F.J. Martín-Reyes, P. Ortega Salvador and A. de la Torre, *Weighted inequalities for one-sided maximal functions*, Trans. Amer. Math. Soc. **319** (1990), 517–534.
- [20] F.J. Martín-Reyes, L. Pick and A. de la Torre,  *$A_\infty^+$  condition*, Can. J. Math. **45** (1993), 1231–1244.
- [21] F.J. Martín-Reyes and A. de la Torre, *Two weight norm inequalities for one-sided fractional maximal operators*, Proc. Amer. Math. Soc. **117** (1992), 483–489.
- [22] F.J. Martín-Reyes and A. de la Torre, *One-sided BMO spaces*, J. London Math. Soc. **49** (1994), 529–542.
- [23] J. Moser, *On Harnack's theorem for elliptic differential equations*, Comm. Pure Appl. Math. **14** (1961) 577–591.
- [24] J. Moser, *A Harnack inequality for parabolic differential equations*, Comm. Pure Appl. Math. **17** (1964), 101–134.
- [25] J. Moser, *Correction to: "A Harnack inequality for parabolic differential equations"*, Comm. Pure Appl. Math. **20** (1967), 231–236.
- [26] S. Ombrosi, *Weak weighted inequalities for a dyadic one-sided maximal function in  $\mathbb{R}^n$* , Proc. Amer. Math. Soc. **133** (2005), 1769–1775.
- [27] O. Saari, *Parabolic BMO and global integrability of supersolutions to doubly nonlinear parabolic equations*, to appear in Rev. Mat. Iberoam, available at arXiv:1408.5760 (2014).
- [28] E. Sawyer, *Weighted inequalities for the one-sided Hardy-Littlewood maximal functions*, Trans. Amer. Math. Soc. **297** (1986), 53–61.
- [29] N.S. Trudinger, *Pointwise estimates and quasilinear parabolic equations*, Comm. Pure Appl. Math. **21** (1968), 205–226.
- [30] V. Vespri, *On the local behaviour of solutions of a certain class of doubly nonlinear parabolic equations*, Manuscripta Math. **75** (1992), 65–80.

DEPARTMENT OF MATHEMATICS, AALTO UNIVERSITY, P.O. BOX 11100, FI-00076 AALTO UNIVERSITY, FINLAND

*E-mail address:* juha.k.kinnunen@aalto.fi, olli.saari@aalto.fi